Available online at www.sciencedirect.com
SCIENCE DIRECT
COGNITION
Cognition 97 (2005) 179–210
www.elsevier.com/locate/COGNIT

Discussion

The evolution of the language faculty: Clarifications and implications

W. Tecumseh Fitch[a,*], Marc D. Hauser[b], Noam Chomsky[c]

[a]University of St Andrews, School of Psychology, St Andrews, KY16 9JP, Scotland, UK
[b]Harvard University, Cambridge, MA, USA
[c]Massachusetts Institute of Technology, Cambridge, MA, USA

Received 5 November 2004; accepted 15 February 2005

Abstract

In this response to Pinker and Jackendoff's critique, we extend our previous framework for discussion of language evolution, clarifying certain distinctions and elaborating on a number of points. In the first half of the paper, we reiterate that profitable research into the biology and evolution of language requires fractionation of "language" into component mechanisms and interfaces, a non-trivial endeavor whose results are unlikely to map onto traditional disciplinary boundaries. Our terminological distinction between FLN and FLB is intended to help clarify misunderstandings and aid interdisciplinary rapprochement. By blurring this distinction, Pinker and Jackendoff mischaracterize our hypothesis 3 which concerns only FLN, not "language" as a whole. Many of their arguments and examples are thus irrelevant to this hypothesis. Their critique of the minimalist program is for the most part equally irrelevant, because very few of the arguments in our original paper were tied to this program; in an online appendix we detail the deep inaccuracies in their characterization of this program. Concerning evolution, we believe that Pinker and Jackendoff's emphasis on the past adaptive history of the language faculty is misplaced. Such questions are unlikely to be resolved empirically due to a lack of relevant data, and invite speculation rather than research. Preoccupation with the issue has retarded progress in the field by diverting research away from empirical questions, many of which can be addressed with comparative data. Moreover, offering an adaptive hypothesis as an alternative to our hypothesis concerning mechanisms is a logical error, as questions of function are independent of those concerning mechanism. The second half of our paper consists of a detailed response to the specific data discussed by Pinker and Jackendoff. Although many of their examples are irrelevant to our original

DOI of original article: 10.1016/j.cognition.2005.04.006
* Corresponding author.
E-mail address: wsf@st-andrews.ac.uk (W.T. Fitch).

0022-2860/$ - see front matter
doi:10.1016/j.cognition.2005.02.005

paper and arguments, we find several areas of substantive disagreement that could be resolved by future empirical research. We conclude that progress in understanding the evolution of language will require much more empirical research, grounded in modern comparative biology, more interdisciplinary collaboration, and much less of the adaptive storytelling and phylogenetic speculation that has traditionally characterized the field.

1. Introduction

In a recent paper, we (Hauser, Chomsky, & Fitch, 2002) (HCF hereafter) offered a framework for research on language evolution, stressing the importance of an empirical, comparative and interdisciplinary approach to this problem, and of distinguishing between several different notions of "language" found in the literature on this subject. In their paper "The Faculty of Language: What's Special about it?" Steven Pinker and Ray Jackendoff (PJ hereafter) present a critique of this paper. In our response, we begin by noting the many areas of agreement between HCF and PJ, among them the need to fractionate language into its component mechanisms, the need for an empirical approach to test hypotheses about these mechanisms, the value of comparative data from diverse animal species for doing so, and the need for collaborative, inter-disciplinary work in this endeavor. However, several distinctions and hypotheses that formed the core of our original paper have been misunderstood by PJ. We first clarify these core ideas before addressing PJ's specific criticisms concerning empirical evidence. Many of their criticisms are based on a mischaracterization of the perspective we outlined; but several substantive areas of disagreement are discussed. This section allows us to clarify certain issues that were left open in HCF, or unmentioned due to space constraints.

1.1. Clarifying the FLB/FLN distinction

One main thrust of PJ's critique results from their blurring the distinction we drew between broad and narrow interpretations of the term "faculty of language." Although PJ endorse this distinction, many of their arguments appear to result directly from a failure to make it themselves, or to perceive where we were making it. We thus start by clarifying this distinction, and its importance.

It rapidly became clear in the conversations leading up to HCF that considerable confusion has resulted from the use of "language" to mean different things. We realized that positions that seemed absurd and incomprehensible, and chasms that seemed unbridgeable, were rendered quite manageable once the misunderstandings were cleared up. For many linguists, "language" delineates an abstract core of computational operations, central to language and probably unique to humans. For many biologists and psychologists, "language" has much more general and various meanings, roughly captured by "the communication system used by human beings." Neither of these explananda are more correct or proper, but statements about one of them may be completely inapplicable to the other. To this end, we denoted "language" in a broad sense, including all of the many mechanisms involved in speech and language, regardless of their

overlap with other cognitive domains or with other species, as the "faculty of language in the broad sense" or FLB. This term is meant to be inclusive, describing all of the capacities that support language independently of whether they are specific to language and uniquely human. Second, given that language as a whole is unique to our species, it seems likely that some subset of the mechanisms of FLB is both unique to humans, and to language itself. We dubbed this subset of mechanisms the faculty of language in the narrow sense (FLN). Although these mechanisms have traditionally been the focus of considerable discussion and debate, they are neither the only, nor necessarily the most, interesting problems for biolinguistic research. The contents of FLN are to be empirically determined, and could possibly be empty, if empirical findings showed that none of the mechanisms involved are uniquely human or unique to language, and that only the way they are integrated is specific to human language. The distinction itself is intended as a terminological aid to interdisciplinary discussion and rapprochement, and obviously does not constitute a testable hypothesis.

We believe that a long history of unproductive debate about language evolution has resulted from a failure to keep this distinction clear, and that PJ, while agreeing with its importance in principle, have not made it in practice. Only this can explain their disagreement with the hypothesis they attribute to HCF at the culmination of their introduction: "that recursion is the only aspect of language that is special to it, that it evolved for functions other than language, and that this nullifies 'the argument from design' that sees language as an adaptation". In any interpretation that equates the last "language" in this sentence with FLB, we not only disagree with this hypothesis (which is not our own), but reject it as extremely implausible. Our focus on the mechanism of recursion in HCF was intended as a plausible, testable hypothesis about a core component of FLB, and likely FLN, not a blanket statement about "language as adaptation." Here, we hope to clarify any possible misunderstanding by exploring these issues in greater detail.

As we argued in HCF, treating "language" as a monolithic whole both confuses discussions of its evolution and blocks the consideration of useful sources of comparative data. A more productive approach begins by unpacking FLB into its myriad component mechanisms. These components include both peripheral mechanisms necessary for the externalization of language, and core linguistic computational/cognitive mechanisms. The proper fractionation of FLB into its components is obviously not trivial or given, and it would be naïve to suppose that the biologically appropriate fractionation will precisely mirror traditional disciplinary subdivisions within linguistics (phonetics, phonology, syntax, semantics, etc.). For example, "phonetics" is traditionally concerned with the sounds of spoken language, while "phonology" concerns more abstract questions involving the mapping between sounds and linguistic structures. Though distinct, both components of language presumably tap some of the same mechanisms. The "phon-" root of both terms reveals their original preoccupation with sound, but language can also be externalized through the visual and manual modality, as in signed languages, raising tricky questions about the precise borders of the sensory-motor component of language. While it would be a mistake to exclude visual/motor mechanisms from FLB, it is not a core component for the vast majority of humans. This is not an issue we attempted to answer in HCF, nor will we do so here. We raise it simply to illustrate the complexity of the issues raised when one attempts to properly fractionate FLB.

In HCF, we offered one potential cut through FLB, explicitly distinguishing the sensory-motor (SM: phonetics/phonology) and conceptual-intentional (CI: semantics/ pragmatics) systems from the computational components of language that have been the traditional focus of study in modern linguistics, including syntax, morphology, a phonological component that interacts with SM systems, and a formal semantic component that interacts with the CI system. We make no claims that this is the only correct way to fractionate FLB, explicitly leaving room for other components (see Figure 1 in HCF). "We make no attempt to be comprehensive in our coverage of relevant or interesting topics and problems" (p. 1570). However, contrary to PJ's suggestion, our framework does not exclude the many important issues that arise in phonology, morphology, or the lexicon. Questions concerning how internal computations relate signal and meaning are explicitly raised in the initial theoretical discussion (p. 1571), and must be, by definition, part of an adequate theory of language.

Something about the faculty of language must be unique in order to explain the differences between humans and other animals—if only the particular combination of mechanisms in FLB. We thus made the further, and independent, terminological proposal to denote that subset of FLB that is both specific to language and to humans as FLN. To repeat a central point in our paper: FLN is composed of those components of the overall faculty of language (FLB) that are both unique to humans and unique to or clearly specialized for language. The contents of FLN are to be empirically determined. Possible outcomes of this empirical endeavor include that ALL components of FLB are shared either with other species, or with other non-linguistic cognitive domains in humans, and only their combination and organization are unique to humans and language. Alternatively, FLN may turn out to include a very rich set of interconnected mechanisms, as assumed in many earlier versions of generative grammar. The only "claims" we make regarding FLN are that (1) in order to avoid confusion, it is important to distinguish it from FLB, and (2) comparative data are necessary, for obvious logical reasons, to decide upon its contents. An equally obvious point is that research on non-linguistic cognitive domains (number, navigation, social intelligence, music, and others) is fundamental to the proper eventual delineation of FLN.

PJ's central complaint with HCF lies with our further hypothesis—stated clearly as such in the paper—that only a relatively compact, but powerful, component of the computational component of language falls into the FLN subset of FLB (Hypothesis 3 of HCF). "We propose in this hypothesis that FLN comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mapping to the interfaces." (p. 1573). The term "FLN" thus served dual duties in HCF. To be precise, we suggest that a significant piece of the linguistic machinery entails recursive operations, and that these recursive operations must interface with SM and CI (and thus include aspects of phonology, formal semantics and the lexicon insofar as they satisfy the uniqueness condition of FLN, as defined). These mappings themselves could be complex (though we do not know) because of conditions imposed by the interfaces. But our hypothesis focuses on a known property of human language that provides its most powerful and unusual signature: discrete infinity. We offered this hypothesis as a starting point for discussion and research, "restricting attention to FLN as just defined but leaving the possibility of a more inclusive definition open to further empirical research" (p. 1571). We do not define FLN as

recursion by theoretical fiat (note, we say "a key component"), which would contradict the aims of our paper, but offer this as a plausible, falsifiable hypothesis worthy of empirical exploration. We hypothesize that "at a minimum, then, FLN includes the capacity of recursion", because this is what virtually all modern approaches to language (including those endorsed by PJ) have agreed upon, *at a minimum*. Whatever else might be necessary for human language, the mechanisms underlying discrete infinity are a critical capability of FLB, and quite plausibly of FLN.

1.2. Biolinguistics and the Minimalist Program

PJ give a long and detailed critique of the Minimalist Program (MP), based on their interpretation of how the minimalist program informed our "overall vision of what language is like." In fact, partly for reasons of space, HCF barely discussed MP. The framework advanced in HCF for the study of language evolution does not rise or fall with the fate of the minimalist program. Indeed, most of our points (e.g. the FLN/FLB distinction, the value of an empirical, hypothesis-testing approach, the importance of comparative data, etc.) apply equally to any of the various flavors of modern generative grammar. Like HCF, our discussion here will be largely non-committal as regards the virtues and faults of the various flavors of generative grammar currently available. The only assumption made in HCF, and here, about syntactic theory is the uncontroversial one that, minimally, it should have a place for recursion. We think researchers in fields outside linguistics should adopt a wait-and-see attitude as these intradisciplinary issues are sorted out. It is certainly not the case that our framework is based on a covert "presumption that the Minimalist Program is ultimately going to be vindicated," and we are quite puzzled by PJ's assertion to this effect.

PJ see minimalism as providing "a rationale" and "motivation" for our hypothesis 3—the only obvious justification for their long and detailed critique of minimalism. This speculation is incorrect. A primary motivation for writing HCF was our recognition of some pervasive confusion that have led to persistent and unnecessary misunderstandings among researchers interested in the biology and evolution of language. Such misunderstanding has polarized debate unnecessarily, has helped to fuel dogmatic and even hostile stances, and has generally acted to block progress in this field, including especially the severing of possible collaborative projects between linguists, psychologists and biologists. It has contributed to a situation in which animal researchers interested in language almost automatically consider themselves anti-linguist, or anti-generative, while some linguists feel justified in being anti-cognitive or anti-evolutionary. The FLN/FLB distinction, we hoped, would help the field to see that there is no incompatibility between the hypotheses that FLB is an adaptation that shares much with animals, and that the mechanism(s) underlying FLN might be quite unique. We further realized that earlier statements that had been interpreted as anti-evolutionary were in fact compatible with contemporary (and perfectly orthodox) neo-Darwinian theory. This realization, not a covert acceptance of minimalist precepts, was the primary motivation for writing HCF, and phrasing our hypotheses as we did.

PJ's comments about MP are thus mostly irrelevant to most of the topics of HCF, and of the current paper, and due to space constraints we are unable to discuss them fully here.

However, lest readers conclude that PJ's characterization of MP is accurate, we devote an online appendix to correcting their many misconceptions about this research program (www.wjh.harvard.edu/~mnkylab). Research in a minimalist framework has made considerably more progress than allowed by PJ, and this research has addressed many of the issues on PJ's list of "ignored" phenomena.

Although we stress the independence of the framework advanced in HCF from the minimalist program, we did suggest and maintain here that a core element of FLN may be structured by considerations of efficient use of the core computational mechanisms of recursion; this is the only place where the discussion in HCF ties in directly to the minimalist program. One implication of this proposal is that much of the complex technology of earlier versions of generative grammar might possibly be eliminated, without losing (and sometimes gaining) empirical coverage (see online appendix). The only practical implication of this additional suggestion is that the study of the evolution of language will be made easier if this proposal turns out to be true, as opposed to more baroque alternative possibilities. If most of FLB builds upon ancient foundations, shared with other animals, we will have a much better chance of understanding these foundations from a genetic and neuroscientific viewpoint, because of the much greater variety of deployable experimental methods. If other aspects of FLB are shared with other cognitive domains (e.g. number, navigation, music, social intelligence) or overarching principles of efficient computation, that will vastly improve our chances of gaining an empirical grip on these mechanisms. The fact that a minimalist FLN would be easier to implement neurally, easier to code genetically, and easier to evolve should hardly be counted as evidence against it. Thus, should the minimalist approach find increasing empirical support, it would be good news for biologists and psychologists.

1.3. What is language "for"?

The hypothesis set laid out in HCF might be interpreted as a set of mutually exclusive alternatives. This would be an error. The hypotheses described represent different perspectives and focus on different targets of analysis. The first two hypotheses concern FLB, and were intended to span the range of currently available hypotheses. In contrast, our hypothesis 3 explicitly concerns only FLN; it is a proposal about what mechanisms are uniquely human. In contrast, PJ emphasize the design features of language "as an adaptation for communication." Crucially, questions of mechanism are distinct from and orthogonal to questions of adaptive function (Tinbergen, 1963), a point clearly distinguished by Hauser (1996). This is not to say that questions at one level do not inform questions at another, as they surely do, but rather, that these are different questions that each require answers. We think PJ's (and others') overly restrictive emphasis on adaptive function is misplaced and counter-productive.

The term "adaptation' conceals a conceptual minefield, long recognized as such by practicing evolutionary biologists (Mayr, 1982; West-Eberhard, 1992; Williams, 1966). Definitions run from diachronic and historical (Gould & Vrba, 1982) to purely synchronic and contemporary (Reeve & Sherman, 1993). Without further specification, the statement that "language is an adaptation" is thus vague enough to have few empirical consequences. In our opinion, there is no question that language evolved, and is very useful to humans for

a variety of reasons. We thought our viewpoint on this issue was made relatively clear in HCF: language evolved, shows signs of adaptive design, and comparative data and interdisciplinary cooperation will be necessary to figure out the details of the evolutionary process. We didn't belabor the point as it seems relatively obvious and is far from a new idea, (e.g. Hauser, 1996; Hewes, 1973; Jackendoff, 2002; Lenneberg, 1967; Lieberman, 1975; Nottebohm, 1976; Pinker & Bloom, 1990). To go further than this requires a more rigorous unpacking of the term "adaptation" with respect to language. We will discuss two aspects in turn, first addressing the current utility of language, then turning to past function(s).

Questions about current utility are (at least in principle) empirically testable. But questions about original function are of a different logical type. It is an unfortunate fact that the two main sources of data to address such historical issues, namely paleontological and comparative, are simply unavailable for behavioral traits unique to one species. This is one reason that some biologists advocate a purely synchronic interpretation of "adaptation" (Reeve & Sherman, 1993). For some behavioral traits there *are* fossil data available to test hypotheses (e.g. we know from fossils that humans adopted a bipedal posture before brain size expansion), and for some linguistic mechanisms there may be relevant comparative data (e.g. for vocal learning). But, considering language as an unfractionated whole, neither type of data is available: "language" does not fossilize and is unique to humans. Thus, from an empirical perspective, there are not and probably never will be data capable of discriminating among the many plausible speculations that have been offered about the original function(s) of language, as for music, mathematical reasoning or a host of other interesting human abilities. We of course do not question PJ's right to speculate along these lines, as long as it remains clear that these speculations are not confused with, or offered as alternatives to, testable hypotheses. However, in our opinion, preoccupation with such questions has served as an impediment to more useful empirical research on the evolution, development, and neurobiological underpinnings of language.

1.3.1. Current utility

Empirically addressing specific hypotheses concerning adaptation requires equally specific hypotheses about function. As we discuss below, "communication" is far too vague to constitute such a hypothesis, and none of the other candidates on offer seem much better. So why argue about them? Consider the analogous question: "What is the brain for?" No one would question the assertion that the brain is an adaptation (in some broad and not particularly helpful sense), but it would seem senseless to demand that neuroscientists agree upon an answer before studying neural function and computation. Even more specific questions like "what is the cerebellum for?" have defied resolution for many decades without blocking detailed and productive empirical research on this neural subsystem. The question "what was the cerebellum originally for?" is hardly even a topic of discussion. This is not to deny the possible utility of adaptive hypotheses in guiding empirical research: suitably specific adaptive hypotheses can serve a useful function in focusing and inspiring empirical research (see Fitch, 2000a; Hauser, 1996 and Section 2.3.2 below). However, there is no need at present for researchers interested in the biology and evolution of language to resolve these issues, or even take a stand on them.

PJ are additionally concerned with the question of what an adaptation is "for". To them, it seems quite obvious that "language is an adaptation for communication." To understand our skepticism about this claim, consider a parallel question: "What is bat echolocation for?" If we interpret the question as one about current utility—"what do today's bats use their echolocation abilities for and how does it contribute to genetic fitness?"—then there are many correct answers. For example, bats use echolocation to find and capture prey (feeding), to navigate, to find mates, and to engage in aerial dogfights with competitors. Bat echolocation is "for" many things, each subserving different aspects of survival and reproduction. Although the majority of pulses are probably used for finding and capturing food, they are simultaneously employed in navigation, and also signal the bat's presence to conspecifics. The question is akin to asking "what is primate vision for?" There are many correct answers, some of them perhaps conflicting, and it would seem odd to stipulate any one of them as "the purpose" of vision or of echolocation. It is hard for us to see the scientific value of framing the question this way.

Returning to language, and distinguishing rigorously between FLB and FLN, the question of "what is FLB for?" clearly has many answers if interpreted in terms of current utility. Today, FLB is used extensively in both communication, and in private thought. The communicative uses can be further subdivided: humans use language in just about every social interaction, including courtship and mating, aggressive interactions with competitors, caring for offspring, sharing information with kin, etc. Thus there can be little doubt that language is useful for communication with other humans, and communication must be one of the primary selective forces that influenced the evolution of FLB. In fact, one of us wrote an entire book on the evolution of communication (Hauser, 1996) based on this observation. But the private uses of language are equally varied and important, including functions like problem-solving, enhancing social intelligence by rehearsing the thoughts of others, memory aids, focusing attention, etc. They seem to extend into almost every domain of thought. Further, many cognitive scientists have speculated, in the Whorfian tradition, that specific details of language may alter thought, creating cross-cultural cognitive differences (Bowerman & Levinson, 2001); if true this would further complicate the relation between FLB and thought. Finally, the phonological component of private speech might help serialize private cognition, focus attention on one train of thought, and increase the capacity of short-term memory. Thus, the phonological component is not necessarily superfluous to "private" uses of language, as PJ assert.

Questions about the specific current utility of FLN are better defined. Accepting for a moment our provisional, tentative assignment to FLN of only recursion and mapping to the interfaces, it seems clear that the current utility of recursive mental operations is not limited to communication. In addition to its clear utility for cognitive functions like interpreting mathematical formulas that are not plausibly adaptations at all, recursive thought would appear to be quite useful in such functions as planning, problem solving, or social cognition that might themselves be adaptations. As an example of how recursion plausibly functions in spatial reasoning and navigation, consider such concepts as ((((the hole) in the tree) in the glade) by the stream) and ask whether there is an obvious limit to such embedding of place concepts within place concepts (… in the forest by the plain between the mountains in the north of the island…). Our proposal that aspects of FLN may function in spatial navigation did not concern dead reckoning or landmark recognition, as

PJ assume, but processes of optimal computation already established in animal navigation, like efficient path integration and no backtracking (e.g. Gallistel & Cramer, 1996). Thus, questions of current utility, while empirically addressable, offer little reason to conclude that either FLB or FLN are useful only in communication.

1.3.2. Functional origins

There is a different way, also interesting, of interpreting the question “what is bat echolocation for?” as a question about functional origins: “What was echolocation for in the first echolocating bats?” This is clearly a different question, and requires different analyses and data if it is to be answered. One potential source of data is the fossil record, but for most traits, paleontological data tell us when a trait appeared, but not why. Analyses of the middle ear in bat fossils suggest that microchiropteran bats evolved echolocation very early in their evolutionary history (Novacek, 1985), but do not tell us what they used it for. And this is exceptional: in the case of behavioral traits like language, even this much fossil data is unavailable. A more promising approach to questions of original function is the comparative method, as emphasized in HCF. If a trait is shared, the first question is whether it evolved independently (“analogy”) in different lineages or is shared by common descent (“homology”). Homologous traits, those shared via common descent, play a central role in comparative biology because they are the key to reconstructing phylogeny (Hall, 1994). However, the existence of homology seriously complicates questions of adaptation, since traits that were adaptive in the common ancestor of some clade are not necessarily so for each species that makes up the clade (Harvey & Pagel, 1991). Thus, while homologous traits are indispensable in systematics, they are not necessarily the traits of choice for the study of function and adaptation.

Organisms can also come to have similar traits through convergent evolution (“analogy”—one form of homoplasy). The discovery and analysis of analogous mechanisms is an equally interesting and important arm of the comparative method, because analogy can provide crucial insights into the adaptive nature of natural selection, independent of inherited details (e.g. Lockwood & Fleagle, 1999; Sanderson & Hufford, 1996). The similarities in body form between dolphins and seals, or in the wings of birds and bats, are independently evolved responses to the physical constraints of swimming and flight respectively. Such similarities provide the surest sign of adaptations to these ways of life, and are crucial in distinguishing adaptation from mere inherited similarity (Gould, 1976). Other types of homoplasy, such as reversions and parallelism, also provide insights into the role of constraints in evolution (Wake, 1991, 1996). Thus, scientists interested in the study of adaptation seek to discover and explore cases of convergence. Returning to bats, if we look at other species which have independently evolved echolocation, echolocating birds appear to use it solely for navigation (Suthers & Hector, 1988), as does a separate bat lineage, the megachiropteran bat *Rousettus* (Möhres, 1956). Echolocating cetaceans certainly use echolocation for navigation, and perceptual abilities to use sound to sense space exist in other mammals, including cats and humans (Griffin, 1958). The one thing common to all of these species is navigation, making the hypothesis that bat echolocation originally functioned in navigation, and was only later specialized for feeding and communication, a plausible one. But while plausible, it is extremely difficult

to test or falsify. The primary value of such hypotheses is to drive comparative work and to lead to more specific, testable hypotheses about mechanistic function.

Regarding the original function of FLB, we advocate a multi-component, multi-functional perspective in HCF, and PJ's commitment to a view of language as a complex adaptation would appear to entail a similar perspective. From this perspective, it seems unproductive to assume, or to seek, a single answer to the question "What was FLB originally for?" An approach to language evolution that fractionates FLB into its component parts is unlikely to come up with the same overarching function for all of these mechanisms. Moreover, if this question is asked, then surely comparative evidence is absolutely critical to a satisfying answer. For example, several researchers have argued that the evolution of major components of FLB were driven by sexual selection (e.g. Darwin, 1871; Miller, 2001; Pinker & Bloom, 1990). For vocal imitation this seems quite plausible (based on comparisons with birdsong or whale song). But, based on comparative data, sexual selection seems unlikely to have driven the communication of complex facts about the world. This aspect of FLB seems more plausibly driven by kin selection or some other non-sexual selection (Fitch, 2004).

Regarding FLN, we wrote: "It is possible, as we discuss below, that key computational capacities evolved for reasons other than communication, but after they proved to have utility in communication, were altered because of constraints imposed at both the periphery … and more central levels." (pp 1569–1570). This statement clearly does not deny a communicative role to FLB or FLN. It does, however, suggest the difficulties inevitably involved in discussions of past function(s) of any unique components of FLB. PJ state that "Chomsky's positive argument that language is not 'for' communication" is based on the use of language as inner speech. In the passage to which they refer (Chomsky 2000, pp. 76–7, not 75), not only is their "positive argument" not proposed, it is explicitly *rejected*. To quote in full: "Furthermore, whatever merit there may be to guesses about selectional processes that might, or might not, have shaped human language, they do not crucially depend on the belief that the system is an outgrowth of some mode of communication. One can devise equally meritorious (that is, equally pointless) tales of the advantage conferred by a series of small mutations that facilitated planning and clarification of thought…—*not that I am proposing this or any other story*. There is a rich record of the unhappy fate of highly plausible stories about what might have happened, once something was learned about what did happen—and in cases where far more is understood" (emphasis added). In discussion elsewhere, the same points about frequency of "inner speech" are adduced to illustrate the pitfalls in trying to determine "function" or "purpose" of a biological system from frequency of use (Chomsky, 2003).

A "positive argument" has been made, most forcefully by prominent biologists, that communicative needs would not have provided "any great selective pressure to produce a system such as language" with its crucial relation to "development of abstract or productive thinking" (Luria, 1974), through its unique property of allowing "infinite combinations of symbols" and therefore "mental creation of possible worlds" (Jacob, 1982). Of course current utility is a poor guide to past function, and it is an open question whether Luria and Jacob are right to question what PJ declare a "truism." But PJ's position on this matter seems obscure. On the one hand, they forcefully deny the Luria-Jacob position. However they also insist on this position, claiming that basic properties of

language derive from prior systems of "recursive thought." There are some concrete proposals about "recursive thought": namely, generative grammars that yield structures at the CI interface (e.g. Heim & Kratzer, 1998; Larson & Segal, 1995). But that cannot be what PJ mean. Perhaps they have in mind a "language of thought," which evolved prior to FLB and includes its basic internal computational mechanisms. But that assumption simply transfers the basic questions of evolution from language to a language of thought, and this new problem cannot even be posed until we are told what the language of thought is. Whatever they may have in mind, PJ's view appears to be that FLB both is and is not an "instrument for expression of thought."

In conclusion, seeking a single adaptive function for "language", treated as a monolithic whole, is more likely to produce confusion and misunderstanding than insight. Treating any complex biological character as if it had a single function is likely to be unproductive at best, if not meaningless. Second, our hypothesis 3 concerns FLN, whatever its contents turn out to be, and not "language;" questions about either its current utility or original function are logically separate from those concerning other components of FLB. Our assertion that FLN is not obviously "for" communication in today's humans, and that there are other equally plausible precursors in past hominids or primates, seems a rather mild one. It is only when FLN is confused with FLB, or current utility conflated with original function, that consideration of this possibility seems unreasonable, or to be "denying a truism."

1.4. Summary

To recap, we take for granted that the large set of complex mechanisms entering into FLB are adaptive in some broad sense, having been shaped by natural selection for, among other things, communication with other humans. We find this idea neither controversial nor particularly helpful in empirical investigations of the biological nature of the language faculty (FLB). Neither Pinker nor Jackendoff have used their theoretical arguments about the adaptive nature of language to fuel empirical work, to our knowledge. However, once FLB is fractionated into component mechanisms (a crucial but difficult process), we enter a realm where specific mechanisms can be empirically interrogated at all levels (mechanistic, developmental, phylogenetic and functional). Each mechanism might have its own separate phylogenetic and functional history, and we expect diverse answers as progress is made in this research program. As a potential example of this process, we offered our hypothesis 3: that FLN is restricted to a simple but powerful recursive mapping capability by definition, unique to humans and unique to the language faculty. This recursive mechanism has some plausible precursors in cognitive domains other than communication. We think these are worthy of more detailed investigation. Thus, while accepting that FLB is an adaptation, we hypothesized that FLN is not an adaptation "for communication." Note that there is absolutely no contradiction between these two statements, as long as the distinction between FLN and FLB is kept clear.

Accepting our terminology, and the necessity of recursion in syntax, there are two ways one could rationally disagree with this hypothesis. First, FLN may include more than the computations subserving recursion and mappings to the interfaces to SM and CI, as we suggest in several places in HCF. If so, our Hypothesis 3 can simply be restated as specific

to the recursive machinery and associated mappings, rather than FLN in full, and all the same considerations will apply. But in either case our hypothesis concerns a specific subset of linguistic mechanisms, not "language" in a broad sense. Second and less trivially, we argued that the mechanisms of recursion and its mappings are simple enough to nullify the adaptationist's "argument from design," a proposition one can question. This more interesting question demands a much better understanding of the neural, developmental and genetic mechanisms underlying recursion and its mappings than currently available. Perhaps the apparent computational simplicity of recursion masks an implementation rich in detailed, fine-tuned shaping of time-hewn parts. But the onus is clearly upon the proponent of this hypothesis to demonstrate this: *a priori*, there is nothing obvious about it. That recursion is useful is obvious; this does not automatically make it an adaptation in the evolutionary sense.

2. What's special: a reexamination of the evidence

Our introductory remarks have clarified why the FLB/FLN distinction is critical for productive discussion of language evolution, and what, precisely, we suggested in hypothesis 3 of HCF. By failing to carry through with this distinction, effectively attributing to us the position that "language" (FLB) is recursion, PJ set up an easily-refuted caricature of our hypothesis. Consider as an illustration their discussion of the genetics of language. PJ cite the data concerning the small but significant changes that have occurred in the human FOXP2 gene as casting "even stronger doubt on the recursion-only hypothesis" because "the possibility that the affected people are impaired only in recursion is a non-starter." PJ conclude that the FOXP2 data refute our hypothesis 3. But as the above discussion will presumably have made clear, these data are irrelevant to our hypothesis. If anything is a candidate for inclusion in FLB but *not* in FLN, it is the FOXP2 gene, a very heavily conserved transcription factor found in all mammals (Enard et al., 2002) and birds (Haesler et al., 2004). Although the precise function of this gene is still not understood, it would be extremely surprising if its function in humans was fundamentally different from that in other mammals. Furthermore, the gene's effect in humans are pleiotropic, including pronounced effects on oro-motor praxis that are independent of its effects on speech articulation. Consequently, this gene is nearly identical in form to a homologous gene in other mammals, and the consequences of its expression are not specific to speech or language. Thus, we find it difficult to understand why PJ would cite the FOXP2 data, detailing an important component of FLB, as refuting our hypothesis about FLN. The same observation applies, *mutatis mutandis*, to many of the observations they present in their critique of our paper. We only work through a few of these here, but encourage the reader to re-evaluate their entire paper in light of these comments.

When the FLB/FLN distinction is maintained, many of the arguments PJ bring to bear simply disappear, at least as criticisms of HCF. In the remainder of this paper we will consider, in turn, the various observations that PJ cite regarding the mechanisms underlying FLB, and organizing our responses using essentially the same numbering system as PJ. We ask for each mechanism whether it might plausibly be a member of the subset constituting FLN. We will conclude that the answer is negative for most of them,

either because clear homologues or analogues exist in nonhuman species, or because the same mechanisms are operative in other nonlinguistic cognitive domains. However, several interesting areas of uncertainty exist, prominently including similarities and differences in vocal learning and lexical acquisition between humans and animals, and regarding the cognitive underpinnings of music and language, which we hope will be a focus of future research.

2.1. Conceptual structure

With regard to the hypothesis that the conceptual structure expressed by language is based upon a foundation shared with other animals, PJ appear to have little disagreement with us. They concur that the rich database of comparative evidence amassed by researchers in comparative psychology and cognitive ethology strongly suggests that we share a significant component of our non-linguistic mental lives with other species. Note that this conclusion has been driven by empirical research and it could quite plausibly have turned out otherwise; indeed for a large portion of the twentieth century the very idea of "animal cognition" was considered laughable by many psychologists. It was only by first discovering clever experimental ways to test for conceptual abilities without language in adults (e.g. mental rotation tasks) and non-verbal infants (e.g. habituation/dishabituation tasks) and second, applying these experimental innovations to animals, that these important empirical advances were possible. This is precisely the approach we are advocating for the study of other aspects of language. Cognitive ethology thus provides evidence of the power of the comparative approach, and an excellent role model for work on language evolution.

In passing we observe that cognitive ethology is still in its infancy and there remains much we do not know about what animals can and cannot do. In a summary of research on mental state attribution, (Hauser, 2000) concluded that the ability to represent others' minds ("theory of mind") was absent in animals, but as the review went to press, Hare and colleagues provided the first solid evidence of this capacity in chimpanzees (Hare, Call, Agnetta, & Tomasello, 2000), with additional evidence piling up soon thereafter. Such rapid progress should make us cautious concerning statements about what is absent in animals, or what is unlearnable without language. For example, PJ cite "ownership" as a concept that is "hard to discern" in animals' naturalistic behavior. But there are many aspects of animal territorial behavior that are difficult to explain without some primitive notion of ownership, such as a "home court advantage" effect that persists even when both contestants are equally familiar with the territory. Detailed experiments on animal "ownership" show how it is influenced by dominance, priority of access, value of resource, and species-specific rules and exceptions (Kummer & Cords, 1992; Kummer, Gotz, & Angst, 1974; Stammbach, 1988). Although one can always find differences, post hoc, between animal and human versions of a behaviour, these data suggest an "ownership" concept in some animals with considerable overlap with our own.

Some of PJ's other suggestions about uniquely human abilities, or concepts "unlearnable without language", are not supported by comparative data. "Multi-part tools" must be an oversight, since the use of dual hammer and anvil stones by nut-cracking chimpanzees is well-attested in at least two chimpanzee field sites (Whiten et al., 1999).

Similarly, we find recent experimental demonstrations of episodic-like memory in jays and rhesus monkeys to be an impressive example of a cognitive representation of time (and space) by non-linguistic animals. Many corvids (crows and jays) hide food for later use (Olson, Kamil, Balda, & Nims, 1995), and sometimes have an extraordinary memory for its location (Balda & Kamil, 1992). Clayton and colleagues have shown that these species also retain a sense of *when* the food was cached, as demonstrated by their failure to attempt retrieval of food that would have spoiled by the time the birds are permitted access. This clever experiment demonstrates that the ability to mark the passing of relatively long periods of time does not require language (Clayton, Bussey, & Dickinson, 2003). Similarly, rhesus monkeys can report when they recall a prior event and when they have forgotten it (Hampton, 2001). Thus our capacity to mark the passage of time, and know what we do or do not remember, is not unique, and may not rely on language at all (of course claiming that the capacity does not *require* language is different from the claim that language does not enhance or somehow modify the capacity).

2.2. Speech perception

Our substantive differences with PJ begin with their discussion of speech (for details see Hauser & Fitch, 2003). We noted above that speech is only one of several viable modalities for linguistic expression (others including signed language and writing), and thus it is important to distinguish the evolution of speech from the evolution of language *per se* (Fitch, 2000a). Nonetheless, speech certainly has pride of place as the default modality when available, and shows many signs of being specially adapted for this role. In their discussion of speech perception, PJ defend the widespread viewpoint that speech perception relies upon speech-specific, uniquely human perceptual mechanisms: that “speech is special” in both of these ways. We agree with PJ, and many other researchers, that “Speech is Special” (SiS) is an interesting, plausible hypothesis. What we rejected in HCF is granting SiS the status of *de facto* null hypothesis. When mechanism X (say, duplex perception, or the McGurk effect, or many others) is discovered in human speech perception, we believe the default assumption should be that this characteristic is also found in animals, *until comparative data are gathered that reject this assumption*. While this stance seems to us the logical one, the opposite stance has historically characterized the SiS debate, and several of PJ’s points suggest that they are prone to the same bias.

2.2.1. “Speech is special” as a default hypothesis

Studies of speech perception and production are a traditional focus of discussions of language evolution, and comparative data have rightfully played a central role in these discussions. We credit Alvin Liberman and his colleagues at Haskins Laboratories with framing the SiS hypothesis strongly and provocatively, and thus spurring a vibrant and important field of research in comparative speech perception. As a result, we know a good deal about animal speech perception. However, Liberman and colleagues used the discovery of categorical perception to argue that speech is special before any relevant comparative data were gathered. At the time, this stance seemed reasonable: the fit between the phenomena of categorical perception and the details of speech (in particular, the shift of category boundaries with place of articulation) seemed too perfect to result

from some general auditory processes. Hindsight being 20/20, we can now see that this fit more likely results from the alignment of speech production to pre-existing nonlinearities in speech perception. Furthermore, the zeitgeist of the time certainly favored the idea that there was nothing special about speech, and the Haskins group was justified in using this and other data to call this into question. However, the discovery of categorical perception in chinchillas, complete with shifting category boundaries, placed SiS in its proper place as a provocative, strong hypothesis to be tested, rather than a default assumption about every new aspect of speech perception (and in this case, rejected). These events should have offered a serious cautionary message to speech researchers from that point on. Unfortunately, in our opinion, this did not occur, and SiS remained, and for the most part remains, the default assumption among speech researchers. It is this stance that we reject.

Given current understanding of neurobiology and comparative psychology, indicating huge overlap in the mechanisms underlying human and animal perception, cognition and action, we suggest that the appropriate default assumption about any newly discovered mechanism is that it is shared between humans and other animals. Human uniqueness is something to be demonstrated (as we do for recursion in Section 2.6), not assumed. In advocating the hypothesis of shared mechanisms, we are expressing a simple commonsense point: don't state that something is not there until you've looked for it. Our comments in HCF about the explanatory landscape in research on language evolution also reflect a relatively conventional general philosophy about the role of strong hypotheses in science. Science progresses by stating and testing falsifiable hypotheses, and the hypotheses most conducive to progress are those that are most readily falsifiable ("strong"), because a falsifiable hypothesis that repeatedly resists falsification is likely to be true. The search for strong, testable hypotheses is of course different from the choice of appropriate null hypotheses. Examples of strong hypotheses are "categorical perception is unique to human speech" or "recursion is unique to human language": either hypotheses can be readily refuted by empirical study. For example, the first hypothesis was rejected both by demonstrations of categorical perception in animals, and of categorical perception of non-speech sounds (Cutting, 1982; Cutting & Rosner, 1974; Rosen & Howell, 1981).

In contrast, the hypothesis that "speech is special" is not strong, because speech requires many component mechanisms, and the demonstration that any one of them is shared with animals does not threaten the hypothesis as a whole. In practice, for each demonstration of a shared mechanism, such as categorical perception, several new mechanisms have risen to take its place that have been postulated as unique to human speech, such as duplex perception or the McGurk effect. Due to the ease of experimentation on humans relative to that on animals, it seems unlikely that this situation will change soon, and we are sympathetic to animal researchers who complain that the SiS hypothesis is a moving target. As PJ say "it would be extraordinarily difficult at present to conduct experiments that fairly compared a primate's ability to a human's." This is precisely why SiS is not a strong hypothesis, nor a suitable default hypothesis in speech research, in our opinion.

We also find the methodological despair implied by this statement puzzling. In our opinion, and many others in the field, it is not "extraordinarily difficult" to compare human and nonhuman abilities. For example, recent comparisons of humans and monkeys have

used the same methods and materials to test aspects of speech perception (e.g. with habituation-dishabituation procedures and no training), finding for example that both species use rhythmic features to discriminate among language groups (Ramus, Hauser, Miller, Morris, & Mehler, 2000); see Section 2.2.2. Such studies provide the basis for more detailed neural work, and it seems parsimonious to assume that when nonhuman primates and human infants show the same capacities with speech, given the same methods and materials and with no training, they are using the same mechanism.

PJ cite data from humans on duplex perception and sinewave speech as evidence "casting doubt on the null hypothesis" of shared mechanisms, but these data are not even relevant to our null hypothesis. Relevant data would involve attempts to demonstrate the phenomena either in animals, or in nonspeech domains, and PJ's statement appears to confuse absence of evidence with evidence of absence. To our knowledge, no one has ever run a study of duplex perception, or the McGurk effect, with nonhuman animals. The simple demonstration of a new phenomenon or illusion in speech perception does not constitute evidence for SiS. As it happens, there *are* relevant data seemingly rejecting the hypothesis that duplex perception is specific to speech (Fowler & Rosenblum, 1990). The same can be said of infants' preference for speech sounds: perhaps monkeys (or dogs, or guinea pigs) prefer speech sounds as well. In the absence of evidence to the contrary, we suggest, it is not just premature but a logical error to see this as casting doubt on the shared mechanism hypothesis. Of course, one can certainly question our assertion that this hypothesis deserves the status of null hypothesis; perhaps from certain theoretical standpoints this is the wrong choice to make. But if one accepts it, as PJ appear to, many of the empirical observations they cite fail to address this hypothesis at all, much less cast doubt upon it.

2.2.2. Comparative studies of animal speech perception

PJ summarize their views about our "shared mechanisms" hypothesis by stating that "the tasks given to monkeys are not comparable to the feats of human speech perception." They base this opinion on several claims with which we disagree. PJ's claim that animal speech research focuses on simple one-bit discriminations accurately summarizes the first twenty years of this research, but ignores progress made in the last five years. Many new studies, cited in HCF, have looked at much more complex discriminations by animals, e.g. the ability to use syllables or vowels to extract the statistical properties of a continuous stream of speech or to extract grammatical rules. PJ's statement that animal speech perception experiments involve extensive operant conditioning is incorrect. While true of earlier work, the field has recently opened to other methods that involve no training, including habituation/dishabituation techniques; and these are precisely the methods used with human infants, that form the core of PJ's hypothesis concerning our evolved specialization. Furthermore, we disagree with PJ's implication that we can only draw weak inferences from training experiments. First, if training is uninformative concerning SiS, this eliminates many studies of human infants that also involve training. Second, training techniques are a powerful tool to determine if a skill can be developed with practice, experience, and attention by an animal. Finally, training studies typically have an initial period of training under reinforcement that is based on a small number of tokens representative of the target category. This training phase is designed, especially for

extremely naïve animals, to teach them the game: that some sounds are rewarded and others are not. Once trained, the animal moves on to the critical generalization phase (often on stimuli that never appear in the training set) in the absence of reinforcement. The generalization phase is critical for inferring the generality of what the animal has learned, under strict experimental control.

PJ think it is unlikely that monkeys have anything like the human capacity "to rapidly distinguish individual words from tens of thousands of distracters despite the absence of acoustic cues for phoneme and word boundaries, while compensating in real time for the distortions introduced by co-articulation and by variations in age, sex, accent, and emotional state of the speaker" (p. 8). But it is precisely the ability to rapidly extract words from distracters, online, and with speaker variation, that recent animal work on statistical learning addresses: cotton-top tamarins can extract the word-like units (trigrams) within a continuous stream of speech (Hauser, Newport, & Aslin, 2001). This finding directly replicates the results of Saffran and colleagues with 8-month-old babies, using the same methods and materials (Saffran, Aslin, & Newport, 1996). Overall, these data refute the claim that animals are incapable of the perceptual feats that humans engage in spontaneously, online, and with distracters.

PJ cite several studies that indicate clear differences between species. The training studies of Sinnott and colleagues are exemplary in that they directly compare the performance of animals and humans on the same or similar perceptual tasks. However, beyond some level of detail it is unclear how to interpret a finding of difference between the performance of an animal and a human subject. Some differences among subjects (of any species) are inevitable, raising the question of which of these differences really make a difference to the language faculty. As a rough guide, we would suggest that only differences between humans and animals that exceed the differences observed among normal human subjects are relevant to questions of human uniqueness. Thus, if macaque monkeys can successfully discriminate /ra/ from /la/, a distinction that poses severe difficulties for adult Japanese speakers, but have slightly different category boundaries than adult English speakers, we do not see this as evidence that the mechanisms underlying perception of this distinction are unique to humans.

To summarize, PJ's review of the literature on animal speech perception is inaccurate and incomplete, ignoring a number of advances that have been made in recent years. Their skepticism about the value of training studies seems ill-founded, and if it were taken at face value it would demand rejection of a significant body of experimental work with human infants (not to mention much of psychophysics). Research on animal speech perception is currently one of the most advanced fields of comparative biolinguistic study, and it is directly relevant to questions of human uniqueness that have dominated discussions of language evolution for many years. A thorough review of this literature (along with results of our own work) led us, in HCF and elsewhere (Hauser & Fitch, 2003), to argue that the traditional claim that "speech is special" needs to be re-evaluated. Claims of human uniqueness should not be made in the absence of at least some relevant animal data, and we currently know of no clear demonstrations of differences between animal and human speech perception relevant to the evolution of language (a possible recent exception is Newport et al., 2004). Thus the safest assumption, at present, is that the mechanisms underlying human speech perception were largely in place before

language evolved, based on either general auditory or vocalization-specific perceptual processes.

2.2.3. Neural data on speech perception

We do not think that the demonstration of double dissociations between speech and environmental sounds provides strong evidence that speech is special. To see why, consider a condition termed alexia without agraphia ("pure alexia"), where after brain injury a patient loses the ability to read but retains the ability to write. Such patients can write individual words, or take dictation, but afterwards are unable to read what they have written. General visual and manual abilities remain intact. Although rare, this syndrome has been repeatedly reported in the neurological literature (Geschwind, 1965; Geschwind & Kaplan, 1962). Similar cases have even been reported with written music (Brust, 1980). From an evolutionary viewpoint, alexia without agraphia provides a cautionary tale. Writing is clearly a cultural development, and alphabetic writing appears to have been invented only once in the history of our species, a few thousand years ago. Given this short timespan, modern human abilities to read and write can hardly be considered adaptations: they are clearly the learned result of interactions between the language faculty (FLB), and more general manual and visual skills, and perhaps other faculties. Thus, finding a discrete brain region or circuit whose destruction impairs reading, but leaves writing intact, is no demonstration that these skills represent genetically determined, functionally specialized adaptations.

We interpret such neural data on writing, along with many other apparently "modular" activities, as providing important insights into the plasticity of mammalian neocortex, rather than evidence of evolutionary specialization to that function. Rather than being hard-wired for particular functions, even sensory cortices appear to be relatively "open" early in development. If one area of cortex that normally processes some data is damaged, another area can take over to serve the same function. For example "rewired" ferrets develop working visual cortex in the temporal lobes (where auditory cortex would normally be found) (Sharma, Angelucci, & Sur, 2000; von Melchner, Pallas, & Sur, 2000), and their visual behaviour is indistinguishable from that of control animals. Similarly, monkeys trained to use a particular finger, or listen for a particular tone, exhibit larger areas of sensory cortex that process these sensory data (Merzenich, Recanzone, Jenkins, Allard, & Nudo, 1989). In humans, musicians show larger cortical areas devoted to piano notes (Pantev et al., 1998), string players have a larger sensory representation in the left (string) hand (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995), and blind Braille readers use their occipital cortex (devoted in sighted people to vision) to process tactile inputs (Sterr et al., 1998a, b). Such examples could be multiplied considerably, but the point should be clear enough. The nature of mammalian neocortex leads us to expect cortical specialization for any task to which an individual devotes considerable time and effort (e.g. reading or writing, or speech vs. environmental sound perception). Discovery of such "modular" specializations is not evidence for (or against) the specific neural circuit being an evolutionary adaptation.

2.2.4. Convergent evolution

PJ's last paragraph of the section on speech perception raises an issue that pervades their discussion of comparative data. PJ describe "comparisons among primates," and

despite some discussion of nonprimate data, relevant non-primate data are often omitted. For instance, Patricia Kuhl hypothesized that the perceptual magnet effect was uniquely human only after finding no evidence for this effect in macaque monkeys (Kuhl, 1991). Later work (Kluender, Lotto, Holt, & Bloedel, 1998), however, demonstrated the most critical component of the effect in an avian species. While PJ are of course perfectly correct in asserting that the perceptual magnet effect is, as far as we know, unique to humans *among primates*, it is unclear why this restriction to primates should be relevant to claims of human uniqueness. When we ask in HCF whether some trait is uniquely human, we are asking whether it is unique among all living forms, not uniquely human among primates, or uniquely human among mammals. The existence of displaced reference in honeybees would mean that displaced reference is not uniquely human, in our usage. The demonstration of recursion in birds would mean that it is not uniquely human, just as surely as the same finding in chimpanzees. The relevant question is the neural, developmental and genetic mechanisms underlying the trait, which may or may not be shared (a question to be addressed empirically). This general point becomes particularly relevant in discussions of convergent evolution in vocal production

2.3. Speech production

2.3.1. Complex vocal imitation

PJ criticize our discussion of both neural and peripheral aspects of vocal production in humans. We noted that the ability to imitate complex sounds vocally ("vocal imitation"), although apparently unique to humans among primates, is found abundantly in songbirds, as well as various distantly-related mammals (e.g. dolphins and seals, Janik & Slater, 1997). Therefore, vocal imitation is not a uniquely human characteristic, and not a component of FLN, despite being a crucial component of FLB. We did not, as PJ suggest, advance this as an argument "against evolutionary adaptation for language in the human lineage." This confuses our exclusion of vocal imitation from FLN with an overall exclusion from language.

The evolution of vocal learning in birds (presumably "for" birdsong) clearly occurred independently of the evolution of vocal learning in humans. But as already discussed, scientists interested in the study of adaptation seek to discover and explore cases of convergence, rather than defining them away. Vocal imitation in birds is quite irrelevant to the question of whether such a mechanism evolved "for" language, as are analogous abilities in whales or in seals. But a discovery of vocal imitation in some relatively unstudied nonhuman primate species would be equally irrelevant. Extant primate species, whether monkeys or apes, are not "ancestral to humans," and the only way to discover characteristics of extinct common ancestors is through application of the comparative method, examining all available data from many species. Abundant data in all primates examined so far indicate a lack of vocal imitation at a level that could support acquisition of a complex lexicon. These data include most importantly our nearest relatives, the great apes (Crockford, Herbinger, Vigilant, & Boesch, 2004; Hayes & Hayes, 1951; Nottebohm, 1976; Studdert-Kennedy, 1983). Thus, the discovery of complex vocal imitation in some new primate species would almost certainly represent an example of convergent evolution, not evidence for homology.

The significance of this observation cuts both ways: the existence of some trait in our nearest living relatives, chimpanzees, does not demonstrate its presence in our last common ancestor (LCA) with chimpanzees. To choose two examples, female chimpanzees in estrous develop extremely prominent sexual swellings. This might seem to indicate the presence of sexual swellings in the LCA, and loss by humans, but additional comparative data belies this conclusion, since the other great apes lack such swellings, as do most other primates. Chimpanzees apparently evolved sexual swellings independently, perhaps in response to their promiscuous mating systems, and humans, rather than chimpanzees, retain the primitive character. In contrast, laryngeal air sacs are present in chimpanzees, but not in humans. All other great apes also possess such air sacs, allowing us to infer that the LCA had air sacs, and humans have lost them in our subsequent evolution. These examples show that a broad comparative database is necessary to draw conclusions about homology vs. analogy, and traits characterizing the LCA.

PJ contend that human vocal imitation is limited to speech production. But human vocal imitation is not specific to language: the use of vocal imitation in music, which PJ cite parenthetically, refutes their contention. All humans can sing (albeit some quite poorly), and even as adults can easily reproduce a novel tune (even if the pitch or key is off). Indeed, the human ability to imitate novel sounds, both vocally and instrumentally, is absolutely central to the cultural transmission of human musical traditions. (Note that the enlargement of the thoracic spinal cord in late *Homo* (MacLarnon & Hewitt, 1999), indicating increased breathing control, is as relevant to song production as to speech). Whether the human music faculty is in some sense parasitic on the language faculty (as Pinker, 1997, has argued) or independent of it (Lerdahl & Jackendoff, 1983), is an interesting open question, but it seems premature to exclude melodic imitation from the domain of vocal imitation, or to assume that human musical skills are non-adaptive, unselected byproducts of language abilities. We are equally unconvinced by PJ's other arguments that vocal imitation is specific to language. That adults are poor at imitating the detailed phonetics of a foreign language should not obscure our adult ability, unusual among mammals, to do a reasonable imitation of complex novel sounds, even those that have no meaning in our own language. Most children enjoy imitating animal sounds, and in many parts of the world, adult hunters are skilled at imitating the sounds of their prey. Finally, the fact that imitative skill is limited by a critical or sensitive period is not evidence that it is language-specific. We conclude from this that (1) vocal imitation is part of FLB, and (2) the question of what it is an adaptation "for" remains open. Even for the much simpler and better studied case of vocal learning in bird song, this question remains open, and again it seems likely that it is "for" many things (Kroodsma & Byers, 1991). We suspect vocal imitation in humans may be similar.

2.3.2. Anatomical issues

In HCF we briefly discussed peripheral anatomical adaptations for vocal production, which we assigned to the SM component of FLB. PJ's discussion again reflects some misconceptions about the comparative method. The paucity of data on animal supralaryngeal articulation makes it premature to conclude that humans have "incomparably" more complex vocal control: the comparisons that exist suggest that animals may be less limited than previously assumed (Fitch, 2000b; Lieberman, 1968). The discovery of

permanently descended larynges in nonhuman animals (Fitch & Reby, 2001; Weissengruber, Forstenpointner, Peters, Kübber-Heiss, & Fitch, 2002) demonstrates that a permanently descended larynx is not uniquely human, as previously believed, and the existence of this trait in speechless nonhumans clearly indicates that it has functions other than increased speech versatility. Size exaggeration is the most plausible candidate explanation, and the factual basis for this hypothesis has been explored in considerable detail (Fitch, 2002; Hauser & Fitch, 2003). These render plausible the hypothesis that the larynx descended originally, in prelinguistic hominids, for purposes of size exaggeration, and that this served as a preadaptation for speech. Of course, like any hypothesis about past function, this one cannot be tested directly. However, current utility can, and the data available indicate that, contrary to PJ's assertion, the evidence for a size exaggeration function, even in modern humans, is quite strong.

Laryngeal descent lowers formants, and signals with lowered formants are perceived as emanating from larger individuals in at least two species (humans and deer:Fitch, 1994; Reby et al. 2005). These data are consistent with, but do not demonstrate, a size exaggeration function for the descended larynx. More telling is the fact that the human larynx undergoes an additional descent, at puberty, and only in males (Fitch & Giedd, 1999; Lieberman, McCarthy, Hiiemae, & Palmer, 2001). Teenage boys do not appear to undergo a corresponding increase in phonetic ability, and indeed girls appear to enjoy a slight advantage in speech ability over boys (Henton, 1992; Hyde & Linn, 1988). The only obvious function for this male-specific pubertal descent of the larynx would appear to be size exaggeration (Fitch & Giedd, 1999; Ohala, 1984): part of a suite of size-exaggerating traits that appear in males at puberty, including broad shoulders and beards. Thus, for adult males, the evidence that this additional laryngeal descent functioned to exaggerate size, as for deer or lions, and does *not* function in any additional speech ability, seems rather strong. Does this mean that the descended larynx, today, is not an adaptation for speech? Of course not. The two hypotheses are independent, and to think otherwise would be to confuse current utility with original function, like saying that bat's wings didn't evolve "for" flight because bats' ancestors used them "for" swimming or walking. We intended no such conclusion. Contrary to PJ's interpretation, nothing in HCF precludes the possibility of a history of selection on and refinement of the supralaryngeal tract for vocal production. Such a claim would be antithetical to some of the most basic principles of evolutionary biology. Note that this also does not mean that the descent of the larynx in human infants is "for" exaggerating size. The fact that this descent starts in babies at age 3 months is a strong argument against this hypothesis, as repeatedly noted (Fitch, 1997; Fitch, 2002; Fitch & Reby, 2001).

The increasingly rich comparative database concerning animal vocal production has interesting implications for the evolution of speech in our species. Specifically, the evidence offers convincing grounds for considering both human speech perception and production to be adaptations that either build upon homologues present in our recent primate ancestors (as comparative primate data indicate) or that have analogues in other more-distantly related species (as with vocal imitation or laryngeal descent). Many of the relevant mechanisms also function in other non-speech domains such as music. For all these reasons, we consider such mechanisms to be part of FLB. They thus do not speak to our hypothesis 3, which concerns FLN. None of these data were offered as "arguments

against evolutionary adaptation for language in the human lineage" as PJ suggest. PJ's discussion simply confuses the issues by failing to maintain the distinction, and attempting to exclude convergent phenomena like vocal imitation from evolutionary consideration.

2.4. Phonology

In HCF we considered phonology as a mapping from narrow syntax to the SM interface. To risk being repetitive, phonology does not represent "a major counterexample to the recursion-only hypothesis" but is irrelevant to the recursion-only hypothesis advanced in HCF. We suggested that the computational resources of recursion and its mapping to SM and CI are the only FLN-specific properties, and that other principles external to language might be responsible for the residual complexity and details of the SM-CI linking. Given our present knowledge, much of phonology is likely part of FLB, not FLN, either because phonological mechanisms are shared with other cognitive domains (notably music and dance), or because the relevant phenomena appear in other species, particularly bird and whale "song". Some regularities in phonology may result from other principles, perhaps organism-independent, that determine computationally efficient mappings from narrow-syntactic objects to the SM interface, a possibility that can be formulated today, but so far resists serious inquiry. Specific aspects of phonology other than those that follow from recursion plus mapping to the interfaces may well be part of FLN—but this possibility should be stated as a hypothesis and tested, not assumed.

Consider music. We agree with the stance urged by (Lerdahl & Jackendoff, 1983) that music and language are independent domains, to be studied with no strong assumptions about their inter-relationships. Nonetheless, the questions that have been framed for language in the generative tradition are precisely the kinds of questions that we should answer for music (Hauser & McDermott, 2003; Jackendoff & Lerdahl, in press). Although it may turn out that music and language have some deep biological connection at the genetic, neural or even evolutionary levels, such a connection can not be assumed *a priori*. Thus we were puzzled by the statement that "major characteristics of phonology are specific to language (or to language and music)" These two statements are not equivalent. To the extent that aspects of phonology are shared in music and language, they are by definition not part of FLN. Only if aspects of music can be shown to be "parasitic" on language, with no independent evolutionary history or mechanistic basis, might these two statements amount to the same thing. Although Pinker has suggested that music represents a non-adaptive by-product of the mechanisms underlying language (Pinker, 1997), this has by no means been demonstrated, and unless it is, the independence assumption of (Lerdahl & Jackendoff, 1983) seems more prudent. In our opinion, the many similarities between linguistic and musical structure provide a fascinating source of potential insight into more general aspects of human cognition, and the many phenomenological overlaps between music and language (e.g. critical period phenomena, congenital and acquired amusias, parallel neural systems, etc.) provide a powerful window into questions of the biological basis for both domains (Koelsch et al., 2004; Peretz & Zatorre, 2005; Trainor & Trehub, 1992). For example, the great variability in exposure to music among humans, with some individuals immersed from an early age and others with very little exposure, provides a powerful tool to explore the degree to which neural specializations for music are

input-dependent. Such extreme environmental differences are difficult or impossible to study in language. Similar empirical leverage has been gained by studying the similarities between birdsong and human music and language (Doupe & Kuhl 1999). The existence of shared mechanisms (e.g. critical periods in song learning) has opened the door to mechanistic analyses in terms of neurobiology and gene expression. As for many other components of FLB, the fact that these features are shared is welcome news to those favoring an empirical approach to their study.

Turning to rhythm, this is a phenomenon clearly shared between language and music (as PJ agree), but it also characterizes dance, another universal human trait that in most cultures is intimately tied to music, and that has received even less formal study than music. From a comparative perspective, there are some intriguing similarities between human music/dance and the dominance displays in our closest cousins, the great apes. Both chimpanzees (Goodall, 1986) and gorillas (Schaller, 1963) perform ostentatious displays which involve a pattern of movements idiosyncratic to the displayer, often with vocal accompaniment, and with a bimanually generated rhythmic component. The "drummed" bimanual rhythm may be obligatory (gorilla chest beating) or optional (chimp buttress "drumming"). Although such bimanual sound making seems quite unusual among animals, this phenomenon has received surprisingly little study (Arcadi, Robert, & Boesch, 1998; Schaller, 1963), and it seems premature to conclude, as PJ do, that rhythm-following is uniquely human, particularly on the dubious authority of "informal observations" in a popular book (Williams, 1967). We agree with (Merker, 2000) that rhythmic entrainment in animals is a capacity sorely in need of empirical exploration.

On the lack of recursion in phonology, we agree with PJ (and most other commentators) that phonology is hierarchical but not recursive (although their statement in Section 4.2 concerning "recursive phonological signals" seems to contradict this). Certainly, syllables cannot be embedded within other syllables indefinitely. However, even in this domain of inquiry, there are many open questions. Syllables are not the only relevant components of phonological structure, and other constituents like intonational phrases seem much better candidates for unlimited self-embedding (Ladd, 1996). While we are by no means convinced that this is the case (nor is Ladd), let us accept for the sake of argument that it is. The discovery of a recursive mechanism in phonology would first raise the empirical questions "is it the same as or different from that in phrasal syntax?" and "is it a reflex of phrasal syntax perhaps modified by conditions imposed at the interface?" Second, given that the phrasal structure of music shows no obvious limit on embedding, we might ask "is phonological recursion the same as or different from that in musical phrases?" or in the phrases of birdsong. If the answer to all of these questions were "same," we would reject our hypothesis 3, possibly concluding that FLN is an empty subset of FLB, with only the *integration* of mechanisms being uniquely human. This is precisely the kind of empirical search and hypothesis testing that we favor.

2.5. Words

Regarding word learning, PJ are correct that we misrepresented the results of (Markson & Bloom, 1997) in saying that children "may use domain-general mechanisms" for learning both words and facts. Properly stated, their results indicate that some mechanisms

underlying word learning are not specific to language. This is one reason that we consider these mechanisms to be part of FLB, not of FLN. Another is that the ability to link novel arbitrary noises to some referent appears to be quite general among vertebrates, present in some form not only in chimpanzees but in parrots, dogs and other species (Kaminski, Call, & Fischer, 2004; Owren, Dieter, Seyfarth, & Cheney, 1993; Pepperberg, 1991). PJ are correct that an empirical demonstration of fast mapping for both facts and words, or by other species, "does not prove they have all their properties in common." Indeed, this hypothesis is virtually irrefutable and can never be "proven" by empirical evidence. Even in the face of new empirical data, one can always cite some new detail or phenomenon of words to rescue the hypothesis. This is precisely the sort of unfalsifiable hypothesis of which psychologists and ethologists interested in language evolution have often, and rightly, complained. Hypotheses like "language is uniquely human" or "word learning is uniquely linguistic" are, in our opinion, too vague and weak to be useful spurs to the kind of empirical research upon which progress depends.

PJ quote our own comments about the vastness of the lexicon as being contradictory to our hypothesis 3. Again, misunderstanding has resulted from a failure to carry through the FLN/FLB distinction. There are many aspects of lexical acquisition that are remarkable, which led us in HCF to the suggestions that the learning capabilities underlying the lexicon might represent an independent, evolved component of language (FLB). Humans have independently evolved many traits (e.g. bipedalism, relative hairlessness, complex tool use, and visual arts) that have no obvious connection to language. Other traits, including fact-learning, vocal imitation, and some musical abilities, appear to overlap with language without being specific to it (hence are part of FLB). There is no contradiction between our hypothesis that the mechanisms underlying word learning, although based on some shared mechanisms and thus part of FLB, have been hypertrophied, streamlined or otherwise specialized to this task in our recent evolutionary history. Nor does this contradict our hypothesis that FLN is limited to the core computational capacities of recursion and mappings to the interfaces. Words have qualities unique to language, just as chess moves have qualities unique to chess and theorem-proving has qualities unique to mathematics. Such observations are not, by themselves, relevant to questions of domain specificity. Nor are they the basis for assigning word-learning to FLN.

In HCF we detailed many of the ways in which we think that both animal vocalizations and the symbols learned by enculturated apes, dolphins and parrots differ from human words. As we stated in HCF, the evidence for reference in animals is weak at best: "Without pursuing the matter here, it appears that many of the elementary properties of words—including those that enter into referentiality—have only weak analogs or homologs in natural animal communication systems, with only slightly better evidence from the training studies with apes and dolphins." (p. 1576). Word meaning may well have characteristics unique to language and distinct from fact learning, or it might not. Work like that of Markson and Bloom sets up exactly the kinds of empirical questions that we argue should take center stage in discussions of the language faculty. Note that such empirical inquiry is completely independent of whether language is or is not "an adaptation for communication."

2.6. Syntax

Syntax clearly plays a significant role in our ability to construct and express new meanings, but at least some of the restrictions and complexities of this process are plausibly inherited from conceptual structure, rather than being part of syntax *per se*. Just as the conceptual structure of objects and events surely influences and constrains the properties of nouns and verbs, it seems plausible to postulate that linguistic devices expressing quantity, tense, aspect or comparison, or other temporal or logical relations, inherit at least some of their structure from the conceptual structure of time, space and logic. The precise locus of such constraints is an active area of current research in linguistics. If there do turn out to be purely syntactic aspects of constituents such as complementizers, auxiliaries, or function words, their existence in other domains (such as music, spatial or social cognition) or in other species, would still require empirical investigation. Such features would not automatically be part of FLN.

Our suggestion that recursion is part of FLN, as defined, is based on the following observations. (1) Recursion is agreed by most modern linguists to be an indispensable core computational ability underlying syntax, and thus language; (2) Despite decades of search, no animal communication system known shows evidence of such recursion, and nor do studies of trained apes, dolphins and parrots; (3) The perceptual data currently available indicate that monkeys cannot even process hierarchical phrase structure, much less recursion; and (4) There are no unambiguous demonstrations of recursion in other human cognitive domains, with the only clear exceptions (mathematical formulas, computer programming) being clearly dependent upon language. Thus, current data justify our placing syntactic recursion in FLN. This assignment would clearly be threatened by a claim of similar recursion in birdsong or the discovery that chimps can process recursive strings, or various other potential empirical findings—all signs of a strong, falsifiable hypothesis. Of course, there's not a lot riding on this, since we don't suggest that only phenomena in FLN are worthy of study. If future empirical progress demonstrates that FLN represents an empty set, so be it. A terminological distinction may well outlive its usefulness. For now, though, our hypothesis 3 seems both plausible and consistent with the available data.

We will discuss PJ's assertion that some of the world's languages might lack "evidence of recursion" only briefly, because this seems to us irrelevant to the questions under discussion. Modern linguistics asks questions about the biological capacity to acquire human language, a set that includes but is not limited to the huge variety that currently exists on our planet. The putative absence of obvious recursion in one of these languages is no more relevant to the human ability to master recursion than the existence of three-vowel languages calls into doubt the human ability to master a five- or ten-vowel language. A Pirahã child raised in a Portuguese, English or Chinese environment will master those languages with the same ease as his or her mother's tongue, just as the same child could learn the recursive embedding principle of parentheses in mathematics, or a computer programming language with recursive structure. In the face of the huge number of human languages that have clausal embedding, the existence of one that does not would in no way alter the explanatory landscape. If anything, this example would seem to add to the grounds for doubting that recursion evolved "for" communication (whatever this means

exactly), if a language is attested that gets along without it. But it surely does not affect the argument that recursion is part of the human language faculty: as Jackendoff (2002) correctly notes, our language faculty provides us with a toolkit for building languages, but not all languages use all the tools.

The inability of cotton-top tamarins to master a phrase-structure grammar (Fitch & Hauser, 2004) is of interest in this discussion primarily as a demonstration of an empirical technique for asking linguistically relevant questions of a nonlinguistic animal. It is clearly too early to conclude that all species are equally hobbled (especially given the paucity of species and methods tested). Nor would we be prepared to draw strong conclusions about innate human abilities until infants or young children have been tested, and until more is known about the neural and psychological basis for the human ability to learn phrase structure (all topics of current investigation). Fitch & Hauser do not even mention recursion in the cited paper, and the generation of limited-depth hierarchical phrase structure was not confused with recursion in that paper (although it was by some commentators on the article). The article does suggest that an inability to perceive and process phrase structure, by any animal, would be a severe impediment to that species' ability to master language. But to the extent that phrase structure is important to music, it would be a correspondingly severe impediment to their mastering a human musical style. If further empirical research shows that no nonhuman species can master a phrase-structure grammar, the hypothesis that animals either lack any recursive mechanisms, or cannot apply them to auditory strings, will be left standing. But if, for example, we discover that songbirds can master phrase-structure grammars, further research will be necessary to determine how they do it, whether this ability involves recursion, whether it is applied across different domains or problems, and whether the mechanism they use is similar or different to that in human beings. The importance of this work is its introduction of an empirical technique capable of addressing these issues, incorporating the formal analysis of language. The empirical technique can be used in a wide range of species, and the hypothesis can be empirically tested and falsified.

2.7. Summary: our view of the evidence

PJ end their summary of the available data with the conclusion "that the empirical case for the recursion-only hypothesis is extremely weak." In our view, most of the data PJ discuss concern mechanisms that are part of FLB by definition, because related mechanisms exist in other species and/or other cognitive domains. These data are thus irrelevant to our hypothesis 3, which concerns FLN. Their conclusion is based on a misreading of our hypothesis postulating that, at a minimum, FLN consists of the core computational capacities of recursion as they appear in narrow syntax and the mappings to the interfaces. This hypothesis is intended as a guide to research; we are interested in the extent to which it is true, but we welcome empirical demonstrations that other mechanisms should be added to FLN. In our view, the most promising data in this regard remain those that we cited in HCF, particularly mechanisms for word learning that can plausibly be hypothesized to constitute human- and language-specific mechanisms. But, in the current

state of knowledge, none of these possibilities can be interpreted as demonstrated, or even rigorously addressed, by empirical data. The comparative data have simply not been collected; due in part to the recency of the specific theoretical proposals and the new methodological advances. Thus, in sharp contrast to PJ, we conclude that our hypothesis 3 is not only plausible, but that no data refuting it currently exist.

3. Conclusion: where do we go from here?

PJ's critique offers a vivid illustration of the problems that can arise if we fail to distinguish the various components of FLB, or if we confuse statements about one component with those concerning another. We doubt that future researchers will need to make a point of distinguishing FLN from FLB at every mention of the word "language," as we have done here. However, keeping the object of discussion clearly delineated would certainly help to avoid future misunderstandings of this sort.

Misunderstandings aside, we feel that numerous areas of agreement, and several areas of substantive disagreement, remain. Agreed is the argument that FLB, as a whole, evolved and functions as a human-specific adaptation with several areas of current utility, one of which is clearly communication with conspecifics. Agreed is the necessity of fractionating FLB into several separate components, each of which might have different evolutionary histories. One of these components, the conceptual-intentional subsystem, is agreed to share major similarities with other vertebrates, meaning that research in comparative cognition and cognitive ethology will play a major role in future discussions of this topic. Topics such as theory of mind in animals are of great interest in this respect. Although it seems likely that some aspects of human conceptual structure are unique to our species, the separate question of whether these are unique to language will require detailed research into nonlinguistic cognition. This highlights a major role for cognitive science in general in this research program, and a need for explicit theoretical models of domains such as spatial reasoning, visual parsing, social cognition and music cognition that can be compared to and contrasted with linguistic theories. Such research is not only interesting in its own right, but will provide insights into the correct fractionation of FLB and the contents of FLN.

We agree with PJ that major aspects of the sensory-motor component of FLB, including minimally vocal imitation and probably some abstract computational mechanisms underlying phonological structure as well, evolved independently in the human lineage. We disagree with PJ's argument that vocal imitation is uniquely human and speech-specific, and thus part of FLN, given the abilities of birds to imitate human speech, and of humans to imitate non-speech sounds (especially song). We think that the many similarities between the mechanisms underlying the human music faculty and phonology provide independent grounds for doubting the language-specificity of these mechanisms. Detailed research into both the structure and function of animal "music" such as bird- or whale-song, and into the neural underpinnings of human music perception and production, will play an important role in resolving these issues.

The most important area of substantive debate appears to center around the computational apparatus underlying language, and especially syntax. PJ argue that

syntax and other formal components of FLB are highly complex adaptations for communication, unique to language, and unique to humans, and thus that FLN is equally complex. They posit that syntax consists of a complex set of independent mechanisms whose interrelations and complexity are the earmarks of adaptation. As we stated in HCF, we agree that this remains a plausible hypothesis, although if true it is very difficult to test. When applied to FLB, as in our hypothesis 2, it is widely agreed upon. But when applied to FLN, its plausibility clearly depends on the complexity of FLN, an issue that is the topic of intense research and far from resolved. If as more comparative data come in, FLN turns out to be empty, as is certainly possible, FLN would not constitute an adaptation by anyone's standard. If FLN is very small, or even limited to the computations underlying recursion and its interfaces as we suggest in our hypothesis 3, its status as an adaptation is open to question; minimally, it will require more evidence than the reliance on complexity and the argument from design. This statement, as we hope is now clear, in no way questions the current utility of language, or the status of the entire FLB as an adaptation. If we take the evolution of language seriously, we must also take seriously the possibility that *any* component of FLB may not constitute an adaptation for language, for communication or "for" anything at all, and this is as true of FLN as any other component. If it turned out that the capacity for recursion resulted from a phase transition in the pattern of neural connectivity that results automatically from increases in neocortex to subcortical tissue ratio, interacting with standard mammalian brain development, this would certainly be an interesting result. If that hypothesis (or a host of similar ones) is ultimately rejected, that will be interesting too, and will represent clear progress. Our main concern is that no one prejudge the issues, and that all plausible hypotheses get a chance to be tested.

In summary, it is clear that there is significant agreement among the various disciplines contributing to an understanding of language evolution, and that the prospects for advancing an empirical science of biolinguistics are promising. Researchers interested in these issues must take seriously the complexity of language and the contributions of modern linguistics to its understanding, while incorporating the impressive gains made in neural, developmental, evolutionary and molecular biology. As it progresses, biolinguistics will help to drive an increasingly rich understanding of human and animal cognition, and to a broad comparative approach to understanding the many shared cognitive mechanisms that are part of the human language faculty. Ultimately, we think it is likely that some *bona fide* components of FLN—mechanisms that are uniquely human and unique to language—will be isolated and will withstand concerted attempts to reject them by empirical research. An understanding of such mechanisms from the genetic, neural and developmental perspectives will illuminate our understanding of our own species. However, the search for such mechanisms should not be the sole topic of interest, but just one of many fascinating questions to be addressed. Ultimately, the question of whether it is uniquely human is far from the most interesting question one can ask about language. We intended the ideas raised in HCF to help accelerate progress in these directions, and we hope that PJ's critique and this response will help clarify the issues and aid progress. And we sincerely hope that research into the biology and evolution of language will not continue to be mired down by the misunderstandings that have so long plagued this field.

Acknowledgements

We thank Stephen Anderson, Andrew Carstairs-McCarthy, Simon Kirby, Robert Ladd, Jacques Mehler, Massimo Piattelli-Palmarini, Michael Studdert-Kennedy and an anonymous reviewer for their comments on earlier versions of this paper.

References

Arcadi, C., Robert, D., & Boesch, C. (1998). Buttress drumming by wild chimpanzees: Temporal patterning, phrase integration into loud calls, and preliminary evidence for individual distinctiveness. *Primates*, *39*, 505–518.

Balda, R. P., & Kamil, A. C. (1992). Long-term spatial memory in Clark's nutcrackers. *Animal Behaviour*, *44*, 761–769.

Bowerman, M., & Levinson, S. (2001). *Language acquisition and conceptual development*. Cambridge: MIT Press.

Brust, J. (1980). Music and language: Musical alexia and agraphia. *Brain*, *103*, 367–392.

Chomsky, N. (2003). Comments on Millikan. In L. Antony, & N. Hornstein (Eds.), *Chomsky and his Critics.* (pp. 308–315): Blackwell.

Clayton, N. S., Bussey, T. J., & Dickinson, A. (2003). Can animals recall the past and plan for the future? *Nature Reviews Neuroscience*, *4*, 685–691.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology*, *110*, 221–243.

Cutting, J. E. (1982). Plucks and bows are categorically perceived, sometimes. *Perception and Psychophysics*, *31*, 462–476.

Cutting, J. E., & Rosner, B. S. (1974). Category boundaries in speech and music. *Perception and Psychophysics*, *16*(3), 564–570.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, *22*, 567–631.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*, 305–307.

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., et al. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, *418*, 869–872.

Fitch, W. T. (1994). *Vocal tract length perception and the evolution of language*. Unpublished PhD thesis, Brown University.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*(2), 1213–1222.

Fitch, W. T. (2000a). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*(7), 258–267.

Fitch, W. T. (2000b). The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals. *Phonetica*, *57*, 205–218.

Fitch, W. T. (2002). Comparative vocal production and the evolution of speech: Reinterpreting the descent of the larynx. In A. Wray (Ed.), *The transition to language* (pp. 21–45). Oxford: Oxford University Press.

Fitch, W. T. (2004). Kin selection and mother tongues: A neglected component in language evolution. In D. K. Oller, & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 275–296). Cambridge, MA: MIT Press.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, *106*(3), 1511–1522.

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, *303*, 377–380.

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society, Biological Sciences*, *268*(1477), 1669–1675.

Fowler, C. A., & Rosenblum, L. D. (1990). Duplex perception: A comparison of monosyllables and slamming doors. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 742–754.

Gallistel, C. R., & Cramer, A. E. (1996). Computations on metric maps in mammals: getting oriented by choosing a multi-destination route. *Journal of Experimental Biology*, *199*, 211–217.

Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, *88*, 585–644.

Geschwind, N., & Kaplan, E. (1962). A human cerebral deconnection syndrome: A preliminary report. *Neurology*, *50*, 1201–1212.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Harvard University Press.

Gould, S. J. (1976). In defense of the analog: A commentary to N. Hotton. In R. B. Masterton, W. Hodos, & H. Jerison (Eds.), *Evolution, brain and behavior: Persistent problems* (pp. 175–179). New York: Wiley.

Gould, S. J., & Vrba, E. S. (1982). Exaptation—A missing term in the science of form. *Paleobiology*, *8*, 4–15.

Griffin, D. R. (1958). *Listening in the Dark*. New Haven, CT: Yale University Press.

Haesler, S., Wada, K., Nshdejan, A., Morrisey, E. E., Lints, T., Jarvis, E. D., et al. (2004). FoxP2 expression in avian vocal learners and non-learners. *Journal of Neuroscience*, *24*, 3164–3175.

Hall, B. K. (Ed.). (1994). *Homology: The hierarchical basis of comparative biology*. San Diego, CA: Academic Press.

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences, USA*, *98*, 5359–5362.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785.

Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford: Oxford University Press.

Hauser, M., Chomsky, N., & Fitch, W. T. (2002). The language faculty: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.

Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

Hauser, M. D. (2000). *Wild minds: What animals really think*. New York: Henry Holt.

Hauser, M. D., & Fitch, W. T. (2003). What are the uniquely human components of the language faculty?. In M. Christiansen, & S. Kirby (Eds.), *Language evolution* (pp. 158–181). Oxford: Oxford University Press.

Hauser, M. D., & McDermott, J. (2003). The evolution of the music faculty: A comparative perspective. *Nature Neuroscience*, *6*, 663–668.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*, 53–64.

Hayes, K. J., & Hayes, C. H. (1951). The intellectual development of a home-raised chimpanzee. *Proceedings of the American Philosophical Society*, *95*, 105–109.

Heim, I., & Kratzer, A. (1998). *Semantics in Generative Grammar*. Oxford: Blackwell.

Henton, C. (1992). The abnormality of male speech. In G. Wolf (Ed.), *New departures in linguistics*. New York: Garland Publishing.

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology*, *14*, 5–24.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, *104*, 53–69.

Jackendoff, R. (2002). *Foundations of language*. New York: Oxford University Press.

Jackendoff, R., Lerdahl, F. (in press). The capacity for music: What is it, and what's special about it? *Cognition*.

Jacob, F. (1982). *The possible and the actual*. New York: Pantheon

Janik, V. M., & Slater, P. B. (1997). Vocal learning in mammals. *Advances in the study of behavior*, *26*, 59–99.

Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for fast mapping. *Science*, *304*, 1682–1683.

Kluender, K. R., Lotto, A. J., Holt, L. L., & Bloedel, S. L. (1998). Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustical Society of America*, *104*(6), 3568–3582.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T. C., & Friederici, A. D. (2004). Music, language, and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*, 511–514.

Kroodsma, D. E., & Byers, B. E. (1991). The function(s) of bird song. *American Zoologist*, *31*, 318–328.
Kuhl, P. (1991). Human adults and human infants show a perceptual magnet effect for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, *50*, 93–107.
Kummer, H., & Cords, M. (1992). Cues of ownership in long-tailed macques, Macaca fascicularis. *Animal Behaviour*, *42*, 529–549.
Kummer, H., Gotz, W., & Angst, W. (1974). Triadic differentiation: An inhibitory process protecting pair bonds in baboons. *Behaviour*, *49*, 62–87.
Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
Larson, R., & Segal, G. (1995). *Knowledge of meaning*. Cambridge, MA: MIT Press.
Lenneberg, E. H. (1967). *Biological foundations of language*. New York, NY: Wiley.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M., & Palmer, J. B. (2001). Ontogeny of postnatal hyoid and larynx descent in humans. *Archives of Oral Biology*, *2001*, 117–128.
Lieberman, P. (1968). Primate vocalization and human linguistic ability. *Journal of the Acoustical Society of America*, *44*(6), 1574–1584.
Lieberman, P. (1975). *On the origins of language*. New York: Macmillan.
Lockwood, C. A., & Fleagle, J. G. (1999). The recognition and evaluation of homoplasy in primate and human evolution. *Yearbook of Physical Anthropology*, *42*, 189–232.
Luria, S. (1974). Comments. In M. Piattelli-Palmarini (Ed.), *A Debate on Bio-Linguistics*. MIT and Centre Royaumont Pour une Science de l'homme.
MacLarnon, A., & Hewitt, G. (1999). The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology*, *109*, 341–363.
Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*, 813–815.
Mayr, E. (1982). How to carry out the adaptationist program. *American Naturalist*, *121*, 324–334.
Merker, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 315–327). Cambridge, MA: The MIT Press.
Merzenich, M. M., Recanzone, G., Jenkins, W. M., Allard, T. T., & Nudo, R. J. (1989). Cortical representational plasticity. In P. Rakic, & W. Singer (Eds.), *Neurobiology of Neocortex* (pp. 41–67). Chichester, NY: Wiley.
Miller, G. F. (2001). *The mating mind: How sexual choice shaped the evolution of human nature*. New York: Doubleday.
Möhres, F. P. (1956). Über die Orientierung der Flughunde. *Zeitschrift für Vergleichende Physiologie*, *38*, 1–29.
Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance: II. Statistical learning of non-adjacent dependencies in a nonhuman primate. *Cognitive Psychology* .
Nottebohm, F. (1976). Vocal tract and brain: A search for evolutionary bottlenecks. *Annals of the New York Academy of Sciences*, *280*, 643–649.
Novacek, M. J. (1985). Evidence for echolocation in the oldest known bats. *Nature*, *315*, 140–141.
Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of Fø of voice. *Phonetica*, *41*, 1–16.
Olson, D. J., Kamil, A. C., Balda, R. P., & Nims, P. J. (1995). Performance of four seed-caching Corvid species in operant tests of nonspatial and spatial memory. *Journal of Comparative Psychology*, *109*, 173–181.
Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (1993). Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology*, *26*, 389–406.
Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*, 811–814.
Pepperberg, I. M. (1991). A communicative approach to animal cognition: A study of conceptual abilities of an African grey parrot. In C. A. Ristau (Ed.), *Cognitive Ethology* (pp. 153–186). Hillsdale, NJ: Lawrence Erlbaum Associates.
Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology.*
Pinker, S. (1997). *How the mind works*. New York: Norton.
Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*, 707–784.

Ramus, F., Hauser, M. D., Miller, C. T., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and cotton-top tamarin monkeys. *Science*, *288*, 349–351.

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society of London, B*, 272(1566), 941–947.

Reeve, H. K., & Sherman, P. (1993). Adaptation and the goals of evolutionary research. *Quarterly Review of Biology*, *68*, 1–32.

Rosen, S., & Howell, P. (1981). Plucks and bows are not categorically perceived. *Perception and Psychophysics*, *30*, 156–168.

Saffran, J., Aslin, D., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Sanderson, M. J., & Hufford, L. (1996). *Homoplasy: The recurrence of similarity in evolution*. San Diego, CA: Academic Press.

Schaller, G. B. (1963). *The mountain gorilla*. Chicago, IL: University of Chicago Press.

Sharma, J., Angelucci, A., & Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature*, *404*(6780), 841–847.

Stammbach, E. (1988). Group responses to specially skilled individuals in a *Macaca fascicularis* group. *Behaviour*, *107*, 241–266.

Sterr, A., Müller, M., Elbert, T., Rockstroh, B., Pantev, C., & Taub, E. (1998a). Changed perceptions in Braille readers. *Nature*, *391*, 134–135.

Sterr, A., Müller, M., Elbert, T., Rockstroh, B., Pantev, C., & Taub, E. (1998b). Perceptual correlates of use-dependent changes in cortical representation of the fingers in blind Braille readers. *Journal of Neuroscience*, *18*(11), 4417–4423.

Studdert-Kennedy, M. (1983). On learning to speak. *Human Neurobiology*, *2*, 191–195.

Suthers, R. A., & Hector, D. H. (1988). Individual variation in vocal tract resonance may assist oilbirds in recognizing echoes of their own sonar clicks. In P. E. Nachtigall, & P. W. B. Moore (Eds.), *Animal sonar: Processes and performances* (pp. 87–91). New York: Plenum Press.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*, 410–433.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology:Human Perception and Performance*, *18*, 394–402.

von Melchner, L., Pallas, S. L., & Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*, *404*(6780), 871–876.

Wake, D. B. (1991). Homoplasy: The result of natural selection, or evidence of design limitations? *American Naturalist*, *138*, 543–567.

Wake, D. B. (1996). Introduction. In M. J. Sanderson, L. Hufford (Eds.), *Homoplasy: The recurrence of similarity in evolution* (pp. xvii–??). San Diego: Academic Press

Weissengruber, G. E., Forstenpointner, G., Peters, G., Kübber-Heiss, A., & Fitch, W. T. (2002). Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*), and domestic cat (*Felis silvestris* f. *catus*). *Journal of Anatomy (London)*, *201*, 195–209.

West-Eberhard, M. J. (1992). Adaptation: Current Usages. In E. F. Keller, & E. A. Lloyd (Eds.), *Keywords in evolutionary biology* (pp. 170–179). Cambridge, MA: Harvard University Press.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (1999). Cultures in chimpanzees. *Nature*, *399*, 682–685.

Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Williams, L. (1967). *The dancing chimpanzee: A study of the origins of primitive music*. New York: Norton.